

03 | 数据的期望值：为什么你坐的飞机总是晚点？

2021-08-02 郭炜

《数据分析思维课》

课程介绍 >



讲述：郭炜

时长 14:37 大小 13.39M

▶

数据给你一双看透本质的眼睛，这里是《数据分析思维课》，我是郭炜。

生活中，你是不是也经常遇到这样的场景，夏天你平时都带着伞，偶尔一天没带伞，结果就正好那天下雨了；不打车时街上到处都是空出租车，但等你需要打车时发现全是满员的；别人的飞机都是正点，你坐飞机永远都是晚点（我自己就是这样）.....

类似的事情还有很多，**总之你越不希望某件事情发生，这件事情往往就会发生**，而在发生之后，有的人就会抱怨说自己运气不好。那真是运气不好吗？这背后有什么原因吗？

其实业内我们经常称这样的“效应”为**墨菲定律**。今天我们就先来说说它的起源，然后再来聊聊它背后反映的数据分析知识。

墨菲的一个玩笑

1949 年，美国的一位航空工程师爱德华·墨菲与美国空军共同研究高速载人火箭“雪橇 MX981”，需要把 16 个精密传感器装在超重实验设备上测试耐压性。可即便是超重实验设备在巨大压力下都变形了，传感器也没有任何的指示。检查后才发现，原来是负责装配的三个同事把这 16 个传感器全都装反了。

对此，墨菲不经意间开了一个玩笑：“如果一件事情有可能出错，让他去做就一定会弄错”。随后的记者招待会上，他的上司斯塔普把这句话称为“**墨菲定律**”，并表述为：“**如果有两种或两种以上的方式去做某件事情，而其中一种选择方式将导致灾难，则必定有人会做出这种选择。**”

随后，人们对墨菲定律又做了更多诠释，比如：

任何事情都不会像它表面上看起来那么简单；

所有任务的完成周期都会比你预计的长；

任何事情只要有出错的可能，就会有极大的概率出错；

如果你预感某件事可能出错，它一定会出错。

背后的数学原理

那么墨菲定律背后的数学原理是什么呢？这里就要引入一个数学概念叫做**期望值**。

什么是期望值？**期望值就是对可能出现的结果以概率为权做加权平均**。举一个简单的例子，你买了一张彩票，有 10% 的概率中 100 元，40% 的概率中 50 元，50% 的概率不中。

那么期望值是 $10\% \times 100 + 40\% \times 50 + 50\% \times 0 = 30$ 元。

这 30 元做什么用呢？买 1 张彩票或者 100 张彩票都看不太出来，但是如果你买 10 万张彩票，那么你中奖的钱很可能会接近 300 万，也就是 30×10 万。所以，这个 30 元的**回报期望值**，衡量了你在足够多的次数下，平均每一次事件的获得的数值。

很多人在数据分析中，往往把“均值”和“平均值”混为一谈。这里我告诉你一个简单区分的方法，那就是用英文来区分：

均值（也叫做期望值）英文是 Mean，它是事前预测的，这个值完全是由概率分布决定，也就是我们前面所说的“对可能出现的结果的概率加权平均”；

平均值叫做 Average，它是事后统计，统计样本值的总和除以样本的个数。

名称	特征	例子
平均值	事后统计，针对既成事实的计算	人均工资、平均身高……
均值（期望值）	事前预测，针对概率分布的预期	预测彩票的投资回报、算牌……



带你了解了二者的概念区分后，我们接着来看二者的关系。我想请你先琢磨下这句话：当样本量 N 趋近无穷大的时候，样本的平均值无限接近数学期望（日常计算时时候相等）。这句话是不是听着很熟？对，这就是上节课我们讲到的大数定律（Law of Large Numbers，LLN）。

简单来讲，**期望就是反映在大数定律下多次执行某件事情之后，得到的一个最可能的收益结果**。例如，我们刚才例子用到的买了 10 万张彩票获利 300 万，平均值（Average）和期望值（或者说均值 Mean）都是 30 元，其实就是利用了大数定律来解释这个现象。

解释墨菲定律

聊完平均值、期望值（均值）和大数定律三者的关系后，我们再回到今天要用数据分析解释的现象：墨菲定律。

人都有一个特殊的心理机制，那就是倾向于记住一些不好的事情。就像航班晚点的概率对每一个人而言都是一样的，但对我来讲，每次飞机晚点的等待就会让我印象深刻，而平时

没有晚点的时候，我的注意力都集中在其他事情上。

现在，我们可以把坏事情的期望定义成 M ，它代表着你记住这个坏事情的概率，坏事情的心理影响定义成 X ，概率为 $R1$ ；再把好事情的期望定义成 N ，它代表着记住这个好事情的概率，好事情的心理影响定义成 Y ，概率为 $R2$ 。根据今天所学，我们就有了如下两个公式。

$$M = X * R1$$

$$N = Y * R2$$

 极客时间

抽象一下，你经常担心一个坏事的发生，比如说事情 A ，我们假设 A 发生时对你产生的心理作用是 X ，事情 A 不发生对你产生的心理作用是 Y 。这个时候显然 X 是大于 Y 的，所以当发生 A 这件事情的时候，你受到的心理影响 X 就比 Y 大。

其实，你在担心一件事情的发生之时，这件事情已经具备了发生的大多数条件。我们假设事件 A 发生的概率是 R ，那么当你担心这件事情发生（ $R1$ ）和不担心这件事情发生（ $R2$ ）的时候，两个概率是不一样的， $R1$ 一定大于 $R2$ 。

这样一来，把我们的讨论代入前面的公式，一个我们担心的坏事的发生期望就像下面这个图所显示一样，对比一下你就会发现，**墨菲定律的原理其实是由我们对于好事情和坏事情的期望值差异造成的**。简单讲，印象深刻再加上担心的时候概率高，自然也就担心什么发生什么了。

坏事情心理影响 $X >$ 好事情心理影响 Y

担心事件发生概率 $R1 >$ 不担心事件发生概率 $R2$

$M = X * R1$, $N = X * R2$, 因此, $M > N$



我们现在用墨菲定律来解释一下这节课开头的情景：

夏天是个多雨的季节，所以你会记得要随身带伞。但恰好有一天你没带伞，并且看天阴沉沉的好像会下雨，于是你十分担心，最后真下雨了，印象自然深刻；

你着急打车的时候一般都是高峰期，再加上你平时高峰期也经常打不到车，于是就更加担心，最后发现全是满员，对打不到车这件事印象深刻；

你在很忙的时候，为了赶时间总选晚上的飞机，前面只要有一个航班晚点了，晚上这一班一定晚点，再加上一旦晚点，你回家就基本半夜了，所以印象尤其深刻。

于是，墨菲定律就产生了。

如何规避墨菲定律

了解了墨菲定律的成因后，你是不是想问，那我们怎么才能避免掉入“墨菲定律”的陷阱呢？

其实影响期望的变量分为两个部分，分别是心理影响大小以及发生概率大小，所以我们可以从这两个方面入手。

对于心理影响来说，我们要做的就是不断调整事情对你心里影响的预期，让它们趋同。特别是遇到坏事情的时候，你可以通过**增加 B 计划等方式，调整预期以降低坏事情发生对你**

的心理影响。

对风险的概率来说，你可以**优化流程，提高自身能力，尽可能减少事件出错的概率。**

这么说可能还是有些抽象，我们来看几个工作与生活中的运用。

“为大概率坚持，为小概率备份”——创业的时候，我们要努力为好的期望（ N ）去坚持，同时考虑为坏的影响（ X ）备份，应该尽力降低坏期望（ $R1$ ）的风险。

“已知的是成本，未知的才是风险”——如果坏的影响（ X ）为已知，那么即使你按照坏事件发生概率（ $R1$ ）100% 来准备资金，这批资金也算是你付出的成本；但如果坏影响（ X ）未知，那么无论坏事件发生概率（ $R1$ ）为多少，都是风险，因为你不知道这个坏事件究竟会造成多大的影响。

“项目风险控制”——项目管理当中有各种风险管理和预防措施，把风险分为很多类，例如静态风险、动态风险、局部风险、整体风险，同时也会把风险应对措施细分为很多类，其实背后的核心是为了去避免墨菲定律的发生，让整体项目在项目经理的期望值下正确运行。

“生活中的风险控制”——我们在生活中，其实也是可以借用这种风险控制的方法论，识别生活中的风险并做好准备，这样才能够在墨菲定律发生的时候不会手忙脚乱。例如提前看看天气预报，查看这趟航班的过往准点率，预估自己乘坐航班情况。在去重要会议的时候，多提前一些时间。这些生活里的小事你或许平时不会太过于在意，但请相信，一旦你把这些小事落到实处，你对生活的掌控力会大大提升。

数据分析解释给你的是现实背后的规律，学以致用才可以让它们发挥最大的价值。

小结

好，总结一下，今天我们通过墨菲定律给你介绍了一个有意思的概念叫期望值（Mean），它是对可能出现的结果的概率加权平均，期望值完全是由概率分布决定。而我们之前讲的平均值一般是指算数平均值，也就是一组数据中所有数据之和再除以数据的个数。某个事情长期不断发生，次数足够多后会达到我们预设的期望值，这就是大数定律。

这几个概念相互依存，又相互不同。其实你可以把平均值、大数定律和期望值这三节课看成一个小整体，对照进行学习。

对平均值来说，你要学会为不同事物去分组，用更细分的数据来看待问题。对于大数定律来说，要成事，其实需要我们不要有赌徒心态，要学会持续投入。而对于期望值来说，平衡预期和未雨绸缪这两个词，希望你能够在生活中灵活运用。

其实，这几个数据分析的概念都告诉我们一个最朴实的道理：没有事情可以一蹴而就（平均值），我们需要努力足够多的次数（大数定律），学会规避风险（期望值）。这样最终在若干年后，企业和个人才能有一份满意的企业 / 个人数据报表。

数据给一双看透本质的眼睛，调整好自己的期望，持续学习持续进步。

课后思考

今天我们把平均值、大数定律和期望值给你串讲了一下，你最近在生活中遇到过墨菲定律的事情么？你觉得怎么做可以减轻和规避这些风险？希望你分享在留言区，我们一起共同成长。

另外，如果你对墨菲定律还有一些其他的认知，也可以写在留言区，我们一块讨论。感谢阅读，如果你觉得这篇文章对你有帮助的话，也欢迎把它分享给更多的朋友。

分享给需要的人，Ta订阅后你可得 **20 元现金奖励**

 赞 3  提建议

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 02 | 大数定律与小数陷阱：生活是随机还是有定数的？

下一篇 04 | 随机对照试验：章鱼保罗真的是“预言帝”么？

精选留言 (13)

 写留言

**芳菲**

2021-08-03

以前一直不喜欢给别人打电话，害怕被拒绝，常常思来想去很久，结果不仅没完成任务还浪费了时间；

后来一方面告诉自己，被拒绝是正常的，根本不用有心理负担；就算被骂一顿，自己也少不了一块肉；

另一方面准备充分，如果对方拒绝自己该怎么应对；...

展开 ▾

作者回复: 是的，长期来做是增大了自己成功的几率。



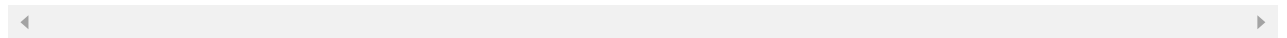
👍 4

**mtfelix**

2021-08-02

记得有个飞行员有个检查清单，这是不是也是对付墨菲定律的法子？那清单法在理论上如何解释？

作者回复: checklist是典型的风险规避法，通过过去的经验和教训，避免风险发生概率r1



👍 2

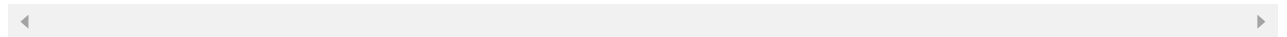
**fly**

2021-08-09

总是丢钥匙，很着急。

多配两个，放在随身带着小包里

作者回复: 一般是办公室会留一个，父母家留一个备份。



👍 1

**Feng**

2021-08-07

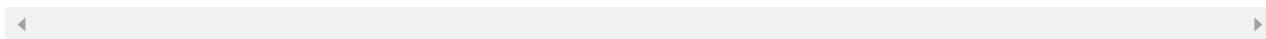
墨菲定律第一次这么明白，用期望值的方法思考清晰了很多，学到了！

**Allen**

2021-08-05

做事前要细心仔细，同时做好心理预期，真碰上事了，能够冷静应对。

作者回复: 一颗平常心！



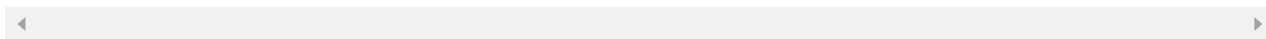
贾诚

2021-08-04

关于墨菲定律在工作中的一点思考，为了达成目标多方面多角度思考问题，哪些因素可以干预影响到结果，如果有别的因素出现，该怎么达成目标，是否有planB，其实很多时候之所以没有取得理想的结果，可能并不是因为工作能力不足，而是因为缺少对结果负责的意识，没有全面考虑到问题

展开 ∨

作者回复: 是的，以终为始。



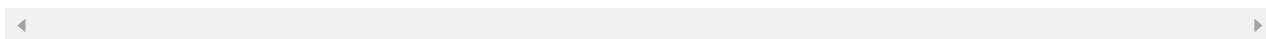
芳菲

2021-08-03

对坏事情发生的结果进行分析，要么提前想好怎么办，要么尝试接受，这样可以降低坏事情发生对心理的影响，从而降低期望值。一旦坏事情发生也就没有那么难受了。

展开 ∨

作者回复: 未雨绸缪，平常心



mmfw

2021-08-03

最近刚换新工作环境，大环境对我们公司不利，要裁员，总是很担心被淘汰试用期过不去，听了老师的大数原理和墨菲定律，学习到应该更长远看问题不纠结眼前得失，墨菲定律也说不要总往坏处想，且要给自己降低被淘汰风险要做到让自己能力增加，这也是降低被淘汰风险。我一直是一个悲观主义者，对即将要发生的事情都经常往坏里想，但过往的事情都结果比较好，我认为是我对待结果的预期很差，结果即使不好，可能是自己已经...

展开 ∨

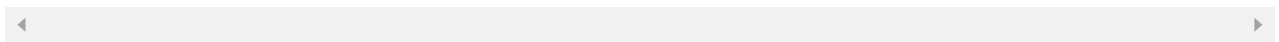
作者回复: 做最坏的打算，做最好的准备，这正是墨菲定律准备方法呀。你做的挺对的，需要提高的是一颗平常心，生活本身就是一种经历，无论是痛苦的还是快乐的，做好准备了就不要过度焦虑，放轻松，体会生活的美好。

**Len**

2021-08-03

请问老师每天几点更新？这都中午12点啦！迫不及待想听老师的新内容了！

编辑回复: 每周一三五零点准时更新哈，也就是说今天（周二）晚上12点就能看到啦~

**进化菌**

2021-08-02

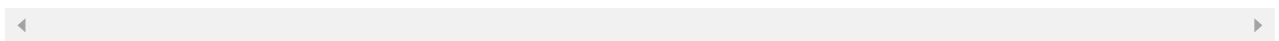
期望值很有意思，越往坏的地方想越容易掉入“墨菲定律”的深坑。

破解墨菲定律，无非就是降低期望值和多做准备，而总在担心是不会改变结果的。

比如最近感觉自己抑郁了，老是在外面看到公司不认识但是知道同公司的人，想上去打个招呼，愣是越想上去越不知道自己为什么不去做，好几次都是最后一犹豫人都走开了.....

展开

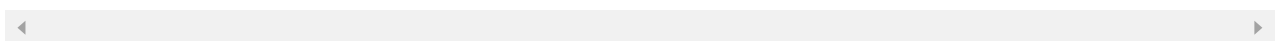
作者回复: 是的，马尔可夫链告诉我们，明天的状态是由今天的状态决定的

**养一条金鱼**

2021-08-02

喜欢一个人，追她时候患得患失，最后往往会追不到也算墨菲定律吧？

作者回复: 是，没有自信，和感觉要失去本身就降低了可能行。

**那一刻**

2021-08-02

墨菲定律，在写代码的时候，会遇到，比如对于某段代码，感觉极小概率出问题，实际上大概率出问题...

感觉可以使用老师提到这句话来解释，“已知的是成本，未知的才是风险”，对于预感到

的错误一定做好处理，另外对于代码逻辑以及相关调用方做好调研，未雨绸缪。 ...
展开 ▾

作者回复: 所以，单元测试，也是增加成本以避免不可知风险的一种呀



微秒

2021-08-02

老师，这里的m和n应该指的是事情对心理影响的期望吧，而不是事情出现的期望？

作者回复: 是的，断句文字稿可能更清楚一些

